Statistical modeling of the influence of a visual distractor on the following eye-fixations

Hélène Queste, Nathalie Guyader, Anne Guérin-Dugué

Statistical modeling of the influence of a visual distractor on the following eye-fixations

Hélène Queste (helene.queste@gipsa-lab.grenoble-inp.fr)[1], Nathalie Guyader[1], Anne Guérin-Dugué[1]

[1]Gipsa-lab, Department Image and Signal, Grenoble, France

Abstract

We examined the influence of a visual distractor appearing during a fixation on the following fixations during natural exploration. It is known that new objects, congruent or incongruent with the scene, appearing during a fixation are fixated more than chance [Brockmole, J. R., & Henderson, J. M. (2008). Prioritizing new objects for eye fixation in real-world scenes: Effects of object-scene consistency. *Vis. Cog., 16*(2-3), 375-390]. In this study, we replicated this result using a Gabor patch for the appearing object, called a distractor because it was artificial and non-related to scenes. Besides, we wanted to quantify its influence on the exploration. A statistical model of the fixation density function was designed to analyze how the exploration was disrupted from and after the onset of the distractor. The model was composed of a linear weighted combination of different maps modeling three independent factors influencing gaze positions. We wondered whether fixation locations observed were rather due to the distractor or the saliency of the scenes. As expected, at the beginning of the exploration, fixation locations were not randomly chosen but influenced by the saliency of the scene and the distractor. The distractor onset strongly influenced fixations and this influence decreased with time.

Method

Stimuli: 156 natural scenes representing a large variety of scenes were used. A distractor, a Gabor patch with a diameter R_d = 2.2°, was added in the scene and appeared about 30 ms after the onset of the third fixation, considered to be the first for the spatial exploration of the scene. Indeed, the first fixation was controlled by a fixation cross to gaze at the periphery of screen, and consequently the second fixation was located around the center of the screen due to central bias. The distractor appeared in a randomly location at 4° from the centre of the scene. Three different presentation times for the distractor were tested: 50 ms (Short Presentation Time: SPT), 210 ms (Medium Presentation Time: MPT) or until the end of the exploration (Long Presentation Time: LPT).

Subjects: 48 participants took part in the experiment (31 female and 17 male; age range: 19 – 32; M=24.03; SD=2.25). 12 observers participated in each distractor condition (SPT, MPT and LPT) and 12 others participated in the Control condition (scenes without distractor). During the Control condition, observers saw the 156 scenes, whereas under the "Distractor" conditions, observers saw only 52 scenes among the 156.

Apparatus and design: Eye movements were recorded using the SR Eyelink 1000 infrared eye tracking system from SR Research Eyelink. Stimuli were presented on a 20-inch ViewSonic CRT monitor at a viewing distance of 57 cm. A trial was a succession of three displays: (1) a white fixation cross presented for 1 s on a mean grey level screen and located on one the four screen diagonals at 5° of eccentricity, (2) a scene displayed for 2.5 s, (3) a grey screen appeared for 1 s. The presentation time of the scene was exactly 2.5 s without the distractor and approximately 2.5 s with the distractor, because we controlled only the time after the third fixation of observers.

Model: From each condition (Control, SPT, MPT, and LPT) and each scene, an experimental saliency map was obtained. We wanted to analyze how the three experimental saliency maps from the conditions with distractor can be explained (1) knowing the experimental saliency map without distractor, (2) taking into account the attractiveness of the distractor and (3) considering all others guiding factors for gaze positions as noise. For each "Distractor" condition, one parametric statistical model was designed assuming additive contributions of these three independent factors [Vincent B.T., Baddeley, R., Correani A., Troscianko, T. & Leonards, U. (2009): Do we look at lights? Using mixture modeling to distinguish between low- and high-level factors in natural image viewing, *Vis. Cog.*, *17*(6-7), 856-879]. The saliency of each scene was modeled with the experimental saliency map (Control condition) by summing Gaussian centered on fixations (sd = 1° to mimic the fovea size and the accuracy of the eye-tracker). The maps were then normalized to 1, to be probability density functions ($Sm(x)$). The attractiveness of the distractor was modeled by a simple Gaussian function ($\mathcal{N}(x;\ \mu(t), \sigma(t))$) where the parameters $\mu(t)$ (mean) and $\sigma(t)$ (sd) were two parameters for the model at each rank fixation t. At last, the third factor was simply modeled by a uniform distribution ($Un(x)$). The additive mixture model gathered these three factors (eq. 1) where $\boldsymbol{x}$ is the 2D fixation positions, $f(x,t)$ the predicted map for one "Distractor" condition (SPT, MPT, or LPT) from experimental gaze positions at each rank t and $\alpha_i(t)$ the estimated weights of each factor $(\alpha_1(t) + \alpha_2(t) + \alpha_3(t) = 1)$:

$$f(x,t) = \alpha_1(t).Sm(x) + \alpha_2(t).\mathcal{N}(x;\ \mu(t),\sigma(t)) + \alpha_3(t).Un(x) \qquad (1)$$

To estimate parameters $(\alpha_i(t), \mu(t), \sigma(t))$ at each rank t, we used the Expectation-Maximization algorithm (EM) [Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. of the Royal Stat. Society, 39*(1), 1-38.].

Results

Fixations that land in the distractor location: We observed that participants predominantly made their fixations in the distractor location for fixation 4, the fixation that immediately followed the distractor onset. As expected, the proportion of fixations that landed in the distractor location decreased as viewing time increased.

Model: Two variants of the model have been tested: one with a mean $\mu(t)$ constant and fixed at the known distractor positions and another with a variable $\mu(t)$ estimated thanks to the EM algorithm. According to the BIC (Bayesian Information Criterion), the best model was for a fixed $\mu(t)$ on the distractor positions. Thus, we only presented the model whose parameters were the weights (Figure 1A) and the standard deviation (Figure 1B). For all the fixations, the average influence of the saliency map (Control) was high (around and more than 60%) and was larger for conditions whose the distractor apparition was short (SPT, MPT) (see the $\alpha_1(t)$ evolution). The contribution of the distractor map increased for fixation 4 ($\alpha_2(4)$), the one immediately following the distractor onset for the three conditions. The uniform contribution representing all the other factors influencing gaze positions remained steadily low except for LPT at the end of the exploration. At the same time, we observed a decrease of $\sigma(t)$ for the fixations 4 and 5 and an increase after, only for SPT and MPT. $\sigma(t)$ represented the distractor attraction: a small $\sigma(t)$ involved that fixations were strongly attracted by the distractor whereas a larger $\sigma(t)$ indicated that the distractor needed to be enlarged to attract fixation locations i.e. the distractor attracted fixations with less precision.

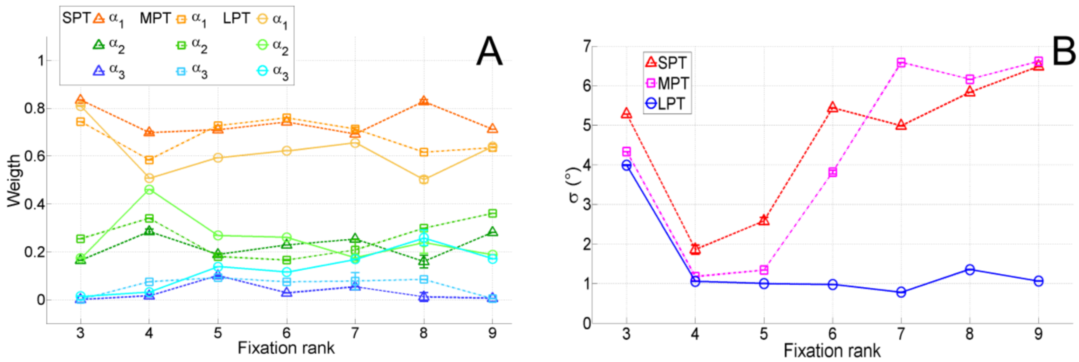

Figure 1: Contributions (weights) of the 3 maps of the additive model (A) and evolution of the $\sigma(t)$ in degrees (B) for the three conditions SPT (triangle), MPT (square) and LPT (circle) as a function of fixation rank.

Conclusion

The proposed model allowed quantifying how fixation locations were influenced by the distractor even when they were not in the distractor location and how this influence continued even when the distractor disappeared from the stimuli. For SPT, the effect was slight after the onset ($\alpha_2(4)$ at 30%, and $\sigma(4)$ less than R_d) and remained small for the fixations following the distractor onset ($\alpha_2(5)$ at 20%, and $\sigma(5)$ around the R_d). It seems that there was no influence of the distractor on late fixations ($\alpha_2(t)$ around 20% and $\sigma(t)$ more than 3 times R_d). For MPT, we observed a larger effect for fixation 4 ($\alpha_2(4)$ at 35%, and $\sigma(4)$ less than R_d). For fixation 5, the distractor attraction was still present but with less precision ($\alpha_2(5)$ around 20% and $\sigma(5)$ around the half of R_d). The most important influence was obtained for LPT: all fixation locations were influenced by the distractor with a good precision due to the non-disappearance of the distractor ($\alpha_2(t)$ higher than 20% with $\alpha_2(4)$ around 50% and $\sigma(t)$ less than R_d).

The results confirmed that fixations were under the direct control of the stimuli and that the distractor onset, which was a transitory phenomenon, disturbed transiently the exploration. The distractor modified fixation locations immediately after its onset but also for several fixations. The disruption was more important when the distractor was presented for a longer period. It seems that the perceptive trace of the distractor influenced the programming of the saccade to the fixation 4 and to a lesser extent, the following saccade to the fixation 5.